Forecasting Research Division
Scientific Paper No. 45

# Neural network classification of convective airmasses for a flood forecasting system

by

G S Pankiewicz

December 1996

Forecasting Research Division
Scientific Paper No. 45

# Neural network classification of convective airmasses for a flood forecasting system

by

G S Pankiewicz

December 1996

Numerical Weather Prediction
Meteorological Office
London Road
Bracknell
Berkshire
RG12 2SZ
United Kingdom

© Crown Copyright 1996

This paper has been accepted for publication in the " International Journal of Remote Sensing" - to be published in 1997.
Permission to quote from this paper should be obtained from the Meteorological Office - Numerical Weather Prediction Division.

# Neural network classification of convective airmasses for a flood forecasting system

by

G.S. PANKIEWICZ

Satellite Image Applications Group (NWP Division),
U.K. Meteorological Office, London Road,
Bracknell, Berkshire, RG12 2SZ, UK

## Abstract

A multilayer perceptron neural network has been trained in order to classify shallow and deep convective airmasses for trial operational use in the GANDOLF thunderstorm flood forecasting system, being developed by the UK Met. Office and the Environment Agency. The airmass classification determines whether to use a precipitation forecast model based on extrapolation techniques, or an object oriented convective precipitation forecast model based on a conceptual model of convective cells. The network was trained on a subset of 2190 Meteosat image samples taken during 1994, each labelled by a meteorologist at the UK Met. Office with the aid of surface pressure analysis charts. The output classes represented by the network include clear land and sea, dynamic cloud and shallow and deep convective cloud. A total of 7 spectral and 2 temporal features were used to discriminate the classes, resulting in a network with a correct classification rate of 80.8±2.2%.

## 1. Introduction

A study is being undertaken at the UK Meteorological Office to examine potential applications for products obtained from meteorological satellite imagery using pattern recognition techniques, particularly in the field of numerical weather prediction. Initially, a number of supervised and unsupervised techniques were considered and compared (Pankiewicz 1995a), including various traditional methods (parallelepiped classifiers, discriminant functions, Gaussian maximum likelihood and dynamic clustering for example), as well as neural network classifiers. The work pointed out that supervised multilayer perceptron neural networks are well suited to such applications, because of their potentially high classification accuracies, their speed in operational use and the fact that no assumptions need be made regarding the statistical distribution of the input features (see also Lee *et al.* 1990, Peak and Tag 1992, Welch *et al.* 1992 and Bankert 1994).

The different applications being considered include cloud field classification, the identification of cloud systems (for example cold fronts, depressions and warm sectors) as well as the direct classification of parameters used in numerical weather prediction models such as relative humidity and mean sea-level pressure. The first two areas overlap considerably, since the identification of a cloud system requires an accurate classification of basic cloud field information (clear, dynamic or convective cloud, for example). To this end, the design of a three-level cloud system classifier was proposed (Pankiewicz 1995a), involving the notions of scale context, since different cloud properties are important at different spatial dimensions. For example, spectral features are more important at the pixel level, whereas textural features lend themselves to regions over which cloud is varying in a significant way within an airmass, perhaps on a scale of tens of kilometres. Likewise, the features that are useful in representing cloud systems, such as shape and size, can be found at scales of several hundred kilometres.

The first-level multilayer perceptron for this three-level classifier was built in 1995 to segment satellite imagery taken from NOAA AVHRR or Meteosat (Pankiewicz 1995b) into basic airmass types (clear land and sea, dynamic cloud and shallow and deep convective cloud), and it was realised that an application for this kind of classifier might be found in a thunderstorm flood forecasting system that was being developed at the UK Met. Office for the Environment Agency.

Known as GANDOLF (Generating Advanced Nowcasts for Deployment in Operational Land-based Flood forecasts), the system was designed to provide automated precipitation forecasts to the Environment Agency and to issue warnings of likely accumulations in regions of interest (Collier *et al.* 1995, Pierce *et al.* 1995). The main sources of error for flood forecasting often occurs during periods of heavy convective precipitation, since the exact timing and location of convective activity is difficult to assess, and most current nowcasting techniques rely on simple extrapolation procedures. A promising new technique which uses conceptual models of convective cells to identify developing and dissipating areas of convective activity has however been undergoing tests at the UK Met. Office (Hand and Conway 1995) and has been incorporated into the GAN-

DOLF system. GANDOLF therefore has the choice of using a precipitation forecast model based on extrapolation (from the UK Met. Office's Nimrod nowcasting system) (Golding 1994) which provides 30 minute and 5km resolution, or a precipitation forecast based on convective cell evolution, providing 10 minute and 2km resolution. The missing link is in determining which forecast to use, and in GANDOLF, a modified version of the airmass classifier now provides one of the main techniques required to make the decision.

## 2. Meteosat sample selection

In order to train the network for use in GANDOLF, 365 Meteosat visible and infrared image pairs were restored from an archive, taken over the course of 1994, one image pair being randomly selected per day in a period from 2 hours after local sunrise to 2 hours before local sunset at the image centre. The archived data originated from the UK Met. Office's Autosat-2 automatic processing system (Allam *et al.* 1990), which obtains both NOAA AVHRR and Meteosat satellite imagery from a ground station at Lasham, Hampshire, and is able to reproject the images into polar stereographic coordinates, and can convert the instrument counts into a range of products, including albedos and infrared brightness temperatures. Meteosat data were selected in this case, since the temporal resolution of 30 minutes provided by Meteosat is more important for GANDOLF than the poorer spatial resolution compared to the 1.1km achievable with AVHRR. The area covered by the imagery is a specific Autosat-2 region (known as area EB), which includes northwest Europe, the North Sea and part of the Atlantic Ocean east of 20°W, and between 45°N and 65°N. In the Autosat-2 system, the nearest neighbour method is used to assign grey-levels to the reprojected pixels, and for the chosen area, the pixel dimensions are 3.84km for Meteosat at 60°N. Examples of two such products taken on 23rd May 1996 (not part of the training data) are given in figures 2 and 3, and will be discussed later in section 4.

Due to a Meteosat data format change in the middle of September 1994, the imagery restored between September 20th and 28th inclusive were considered to be of poor quality for sample selection. Instead, extra samples were taken for 5 days preceding and 4 days following those dates. Otherwise, a total of 6 samples per day were selected to provide a sufficiently large training set containing a range of cloud types which was aimed at reflecting the true distribution found in the imagery. The samples were selected by Clive Pierce, a meteorologist at the UK Met. Office, using surface analyses (including the positions of significant frontal features) to verify airmass information.

Each sample consisted of a box of 17x17 pixels, to ensure sufficient statistics and provide information on the airmass as well as the cloud type. A total of 2190 samples were selected and labelled according to one of four classes: clear land or sea, dynamic cloud and shallow and deep convective cloud. Dynamic cloud is taken here to include a large range of cloud type, such as frontal cirrus, thin cirrus, altostratus, nimbostratus, stratus, fog and haar. Shallow convection includes most forms of stratocumulus and cumulus (cirrus overlying stratocumulus was generally labelled to be shallow convection,

since the classification for GANDOLF is more relevant for cloud at lower altitudes). Deep convection was taken to include cumulonimbus clouds and mesoscale convective systems. Finally, the selected samples were quality checked to ensure that they did not overlap other samples and were correctly classified.

For the purposes of training, the samples were split into two groups: a training set containing two-thirds of the samples, and a testing set with the remaining third. Samples were randomly selected from all four classes into these groups to provide the distributions shown in table 1. It is clear from this table that the number of samples in each class are not the same; they are however a measure of the *a priori* distributions of the classes actually encountered in the imagery, which can be a significant factor in terms of classification accuracy when training neural networks (Foody *et al.* 1995).

| | CL | DY | SH | DE | Total |
|---|---|---|---|---|---|
| Training | 271 | 677 | 446 | 66 | 1460 |
| Testing | 161 | 309 | 206 | 54 | 730 |
| Total | 432 | 986 | 652 | 120 | 2190 |

Table 1: Distribution of training and testing samples by class from the total set of 2190 samples. (CL = clear, DY = dynamic, SH = shallow convective, DE = deep convective).

## 3. Training the neural network

A conventional multilayer perceptron neural network was used for the GANDOLF classifier, with training based on the backpropagation learning algorithm (Rumelhart *et al.* 1986). Here, an error function is calculated at each output node, and is backpropagated through the network in order to update the interconnecting weights between each node. In this way, the error function decreases to a minimum after repeated presentation of training samples, each presentation being referred to as an epoch. For a network with two active layers $j$ and $k$, the error function at the output layer $k$ is:

$$\delta_k = \sigma a_k(1 - a_k)(o_k - a_k) \tag{1}$$

where $\sigma$ measures the spread of the thresholding function, $a_k$ is the estimated network output value for each class and $o_k$ is the true output value for each class. For the preceding layer, the error function is:

$$\delta_j = \sigma a_j(1 - a_j) \sum_k (\delta_k w_{jk}) \tag{2}$$

In the GANDOLF network, the weights between layers $k$ and $j$ were updated with the following simple gradient descent algorithm, the technique used to search for the

minimum value of the error function:

$$w_{jk}(t+1) = w_{jk}(t) + \rho\,\delta_k a_j \tag{3}$$

$\rho$ being the rate of adaption or learning rate. For weights between $i$ and $j$, the same equation was used with $j$ replaced by $i$, $k$ replaced by $j$, and so on.

Numerous parameters affect the performance of such a network, and those which were varied in this application included the number, type and normalisation of features calculated in the image samples, the number of hidden layers and nodes, the range of output values $o_k$ used in training, the spread of the sigmoid thresholding function $\sigma$, the number of training samples presented at each epoch and the rate of adaption of weights $\rho$.

Selecting the features to use in the network is itself an important pattern recognition process, and for this study, 7 simple spectral features were used, based on proven cloud classification features (Gu *et al.* 1991, MacLaren *et al.* 1994 and Bankert 1994). All features were normalised in the range 0 to 1 and the output values were set to either 0 or 1 according to the class of the training sample. Large weight values were not encountered in practice (sometimes a symptom of setting feature and output values in this range), since very few feature values occurred near the 0 or 1 limits. This simple method of normalisation was also found to provide fewer misclassifications than using a zero mean and unit variance normalisation for example. The features used, together with their normalisation factors are shown in table 2, and the means and standard deviations of the features for each class are displayed in figure 1.

| Channel | Feature | Normalisation |
|---|---|---|
| Visible | Minimum | 255.0 |
| Visible | Maximum | 255.0 |
| Visible | Standard deviation | 64.0 |
| Infrared | Maximum | 255.0 |
| Infrared | Minimum/Maximum) | 1.0 |
| Infrared | Standard deviation | 64.0 |
| Visible | Mean Roberts gradient | 64.0 |
| - | Month of year | 12.0 |
| - | Hour of day | 24.0 |

Table 2: The nine features used in the cloud classifier together with their normalisation factors.

A word of caution should be given about the use of maximum and minimum values in this network. Normally, visible grey-scales are considered to have black representing low albedo and white representing high albedo, so that clear sea looks dark and high clouds look bright. For the chosen Autosat-2 Meteosat products, visible grey-scales

range from a value of 4 to 251 representing a range of 100% to 0% in albedo respectively (the reverse of that normally expected). Likewise, infrared grey-scales also range from 4 to 251 and represent a range of 198K to 308K in brightness temperature respectively. In this work, the original grey-scale values have been kept so that clear samples show a large mean of about 0.80 (see figure 1), corresponding to an albedo of about 20%. In the usual view of images with white cloud and dark sea, these feature values are related in the opposite sense (minima becoming maxima).

From figure 1, it can be seen that the 7 spectral features act as very simple yet relatively effective discriminators for the chosen classes. This is to be expected however, since albedo provides a measure of the cloud liquid water in the sample, brightness temperature measures the cloud-top temperature, and standard deviations and gradients give an indication of any cloud structure within the sample, without having to resort to textural features. Additionally, 2 simple temporal features (month of year and hour of day) were used to help discriminate dynamic and convective cloud. (A slightly higher monthly mean was found due to thicker cloud cover associated with frontal systems in winter and slightly higher daily means of convective cloud were found, which typically occur later in the day.) No second order textural statistics were considered in this network, for the reasons discussed in the introduction.

One hidden layer was found to be sufficient for the classifier, particularly as a second hidden layer requires a much longer training time. For the 9 features used (as well as a bias node set to 1), and the 4 output classes, 12 hidden nodes were found to provide fewer misclassifications than for smaller or larger numbers of hidden nodes. The result was therefore a 10:12:4 network architecture.

A sigmoidal thresholding function was used to control the effect of very large or small feature values, in which the output or activity on a node at layer $k$ is given by:

$$a_k = \frac{1}{1 + e^{-\sigma I_k}} \tag{4}$$

where $I_k$ is the net input to the node from the previous layer $j$. The value of the spread $\sigma$ was set equal to 1.0 throughout, as changing $\sigma$ was found to have no significant decrease in the number of misclassifications.

A rule-of-thumb for the number of training samples to use per class is 10 to 20 times the number of classes (Tovinkere *et al.* 1993). However, it has been pointed out that the number of training samples used can affect network accuracy, such that the number should ideally represent the *a priori* distribution of samples in the imagery (Foody *et al.* 1995). In practice, the number of samples used in the GANDOLF network were based on classification accuracies obtained from a number of trial networks, the optimum set being found to contain 40, 160, 120 and 40 samples (drawn randomly from the full set of 1460 samples) for each of the clear, dynamic, shallow and deep convective classes respectively. This was sufficient to combine effects of the *a priori* distribution, the rule-of-thumb concerning the minimum number of training samples needed and the desire to keep to a sensible amount of training time.

To avoid overfitting the network solution to the features in the training samples, a method was used to determine when to stop training. The approach taken was simple and pragmatic, so as to obtain a low misclassification rate without spending too long looking for the optimal solution. This consisted of setting the rate of adaption of weights to a deliberately large value of $\rho = 0.9$, to provide a gradient descent that would cover a large portion of the error surface, but that might overshoot the global miminum. By checking the network behaviour every 5 epochs in terms of the change in mean square error at the output layer, training was stopped when this value was less than $10^{-5}$ for 10 consecutive epochs, or when a maximum number of 200 epochs had passed. After training, the weights were examined every 5 epochs by testing the network with 200 samples (50 per class) drawn randomly from the test set of 730. In this way, training would take a maximum of 20 minutes on a DEC $\alpha$ 3000-600 workstation and the results could be tested to see if the network had reached a solution good enough for use in GANDOLF. When a good solution was obtained, an estimate of the correct classification rate was made by calculating the average from 10 such tests. A large standard deviation in the correct classification rate indicates the solution to have become stuck in a local minimum (MacLaren *et al.* 1994), as does the contents of the associated confusion matrix, in which the classes estimated by the network are tabulated against the true classes for a given number of samples (see below).

| | Est CL | Est DY | Est SH | Est DE | Total |
|---|---|---|---|---|---|
| True CL | 49 | 1 | 0 | 0 | 50 |
| True DY | 7 | 30 | 8 | 5 | 50 |
| True SH | 0 | 2 | 42 | 6 | 50 |
| True DE | 1 | 4 | 4 | 41 | 50 |
| Total | 57 | 37 | 54 | 52 | 200 |

Table 3: A typical confusion matrix from the GANDOLF network with a correct classification rate of 81.0%. True class numbers are given in the rows, and class numbers estimated from the network are given in the columns.

| | Est CL | Est DY | Est SH | Est DE | Total |
|---|---|---|---|---|---|
| True CL | 158 | 3 | 0 | 0 | 161 |
| True DY | 56 | 185 | 50 | 18 | 309 |
| True SH | 2 | 8 | 167 | 29 | 206 |
| True DE | 1 | 4 | 4 | 45 | 54 |
| Total | 217 | 200 | 221 | 92 | 730 |

Table 4: The confusion matrix for all of the test data used to examine the GANDOLF network with a correct classification rate of 76.0%. True class numbers are given in the rows, and class numbers estimated from the network are given in the columns.

With the features, parameters and training sets described above, the best correct classification rate obtained was 80.8±2.2% when examined in this way, or 76.0% when tested with all 730 test samples. Tables 3 and 4 show a typical confusion matrix and the confusion matrix for the full test set used for this network. A major source of error (15.4%) arises from the difficulty in classifying dynamic cloud. For example, some cases of thin cirrus were misclassified as clear and some cases of broken stratus were misclassified as shallow convection. This last type of error may have resulted from a desire during selection of the training samples to label stratocumulus overlain by cirrus as shallow convection, and this may have had the effect of training the network to overestimate dynamic samples as shallow convection. Previous work has shown that it is possible to classify thin cirrus to around 80% accuracy (Bankert 1994, MacLaren *et al.* 1994), similar to the overall accuracy of the network described here, but using *textural* features. It seems likely that the use of texture would have improved both of these sources of error, but would have compromised speed of operation. However, it became apparent that the known sources of misclassification would not be a major problem when applied for use in GANDOLF.

## 4. Current use of the classifier in GANDOLF

As remarked in the introduction, the idea of using a neural network classifier in GANDOLF is to distinguish convective from dynamic airmasses, in the sense of the cloud described earlier, in order to decide whether to use a precipitation forecast model based on extrapolation, or one based on an object oriented forecast model incorporating a conceptual model of the evolution of convective cells. Some simple logic is used in GANDOLF to determine whether the airmass is predominantly dynamic or convective, based on the number of differently classified samples.

First however, the classifier is used on samples of 4x4 pixels instead of the original training size of 17x17 pixels. This loses little information in terms of the classification accuracy, except that the centres of deep convective cells tend to become classified as dynamic cloud, which is strictly true given that they consist of cirrus tops, which were earlier put into the dynamic cloud group. The advantage of this procedure is that it speeds up classification times dramatically, enabling the complete image to be classified in 90 seconds on a DEC $\alpha$ 3000-600 workstation. The classified image is then reprojected into the same area used by the UK Met. Office's Nimrod forecasting system, which covers roughly the central two-thirds of the Autosat-2 EB product area. The sampling of 4x4 pixels (15x15km at 60°N) used in the classifier also coincides with the grid squares used in the Nimrod forecast model.

Examples of a pair of visible and infrared Meteosat EB area images taken on 23rd May 1996 at 10:30 UT are given in figures 2 and 3. Note that the grey-scales have been flipped to show cloud as brighter regions and clear or warm areas as dark, compared to the grey-scales described in section 3 for feature calculation. Figure 4 shows the corresponding classified image for the smaller Nimrod area, obtained with the GANDOLF neural network. In this particular example, a frontal wave can be seen approaching the

southern half of the UK, with an occluded front to the north and northeast of Scotland. Cooler unstable air is advancing from the northwest, providing shallow convection, but also a little deep convection a few hundred kilometres west of Scotland. A clear slot is evident over the western half of the North Sea, as is a narrow band of cirrus stretching from the northeast coast of England out to the Netherlands.

Within the classified area, individual class samples are counted over a smaller region, previously specified by the Environment Agency, covering approximately 2.5°W to 0.5°E and 51.0°N to 52.5°N, and containing about 140 samples. The following logic is then applied to these classified samples:

**At night**, both convective and dynamic precipitation fields are taken from the Nimrod forecasting system.

**During the day**, if all of the classified samples within the region of interest display shallow or deep convection, the object oriented precipitation forecast is used. Alternatively, if the number of samples classified as shallow and deep convection are greater than the number of samples classified as dynamic, *and* if the average convective available potential energy (CAPE) in the region is greater than $10 \mathrm{Jkg}^{-1}$, the object oriented forecast is used. Otherwise, the extraploation based Nimrod forecast is used. Estimates of CAPE are obtained from the Nimrod forecast model and have been discussed elsewhere (Pierce *et al.* 1995).

In cases where dynamic cloud is being misclassified as clear or shallow convection (see the discussion in section 3), but the airmass is predominantly dynamic, GANDOLF will use the independent CAPE information to ensure that the correct airmass has been identified. In addition, misclassifications between shallow and deep convection are ignored in this scheme, so the ability of the network to detect convection increases from an average of 82.2% (table 4) to 94.2%. The use of the neural network together with CAPE in the way described therefore provides a very accurate method for discriminating airmasses, much greater than the quoted accuracy for the network alone of 80.8%, and it is believed that this arrangement is sufficient for operational use in GANDOLF.

The trained neural network code was installed in the GANDOLF nowcasting system in April 1996. Since the beginning of May, the system has been undergoing trial operational use with precipitation forecasts and warnings being issued to the Environment Agency, and it is expected to continue until the end of October 1996, after which it is hoped that the system will become fully operational.

## 5. Summary and future work

A neural network has been described for use in a meteorological nowcasting system to distinguish convective and dynamic airmasses. The philosophy has been to use very simple techniques in order to produce an efficient yet robust and operationally fast

classifier. These include a basic multilayer perceptron with an architecture of 10:12:4 nodes, a gradient descent algorithm using only the error function and a high rate of adaption, and 9 features based on the maximum, minimum, standard deviation and Roberts gradient in some of the visible and infrared Meteosat channels, as well as the month of year and hour of day as two temporal features.

Training and testing on independent samples taken from a total of 2190 Meteosat samples labelled by a meteorologist resulted in a network with a correct classification rate of 80.8±2.2%. The network has now been installed in a trial operational version of the GANDOLF nowcasting system, to provide precipitation forecasts and issue warnings to the Environment Agency. To this end, GANDOLF uses the network to discriminate clear, dynamic and convective regions over the southern part of the UK and incorporates convective available potential energy to assess the predominant airmass, providing a degree of accuracy much higher than would be available from the 4 class network alone.

The work is part of ongoing research at the UK Met. Office to look at the use of pattern recognition techniques for identifying cloud and cloud systems in satellite imagery for a number of meteorological applications, but in particular, numerical weather prediction. The next stage in the construction of the three-level system described in the introduction will be to segment the classified imagery into regions suitable for the calculation of textural features.

## Acknowledgements

The author would like to thank Clive Pierce of the UK Met. Office for his time and patience in helping to label over 2000 Meteosat image samples, as well as for his effort in incorporating the classifier into the GANDOLF system. Acknowledgements are due to Dr. Bryan Conway and Keith Whyte of the Satellite Image Applications Group at the UK Met. Office for helpful discussions throughout.

## References

Allam, R.J., Dibben, P., Holpin, G., Smith, F. and Standley, A.P., 1990, The use of satellite imagery products for operational meteorology. *Satellite Meteorology and Oceanography, Proceedings of the 5th Conference held in London, UK, on 3-7 September 1990*, (Boston: American Meteorological Society), pp.401-403.

Bankert, R.L., 1994, Cloud classification of AVHRR imagery in maritime regions using a probabilistic neural network. *Journal of Applied Meteorology*, **33**, 909-918.

Collier, C.G., Hardaker, P.J. and Pierce, C.E., 1995, The GANDOLF system: An automatic procedure for forecasting convection. *27th Conference on Radar Meteorology held in Vail, Colorado, on 9-13 October 1995*, (Boston: American Meteorological Society), pp.606-608.

Foody, G.M., McCulloch, M.B. and Yates, W.B., 1995, The effect of training

set size and composition on artificial neural network classification. *International Journal of Remote Sensing*, **16**, 1707-1723.

GOLDING, B.W., 1994, Nimrod = FRONTIERS+NWP. *Mercury, (Newsletter of the UK Meteorological Office*, **92**, 8.

GU, Z.Q., DUNCAN, C.N., GRANT, P.M., COWAN, C.F.N., RENSHAW, E. AND MUGGLESTONE, M.A., 1991, Textural and spectral features as an aid to cloud classification. *International Journal of Remote Sensing*, **12**, 953-968.

HAND, W.H. AND CONWAY, B.C., 1995, An object-oriented approach to nowcasting showers. *Weather and Forecasting*, **10**, 327-341.

LEE, J., WEGER, R.C., SENGUPTA, S.K. AND WELCH, R.M., 1990, A neural network approach to cloud classification. *IEEE Transactions on Geoscience and Remote Sensing*, **28**, 846-855.

MACLAREN I., WALDER P. AND REID C., 1994, Neural networks and cloud classification. *Meteosat Scientific Users' Conference, Proceedings of the 10th EUMETSAT Conference held in Cascais, Portugal, on 5-9 September 1994*, EUM P15 (Darmstadt: EUMETSAT), pp.541-547.

PANKIEWICZ G.S., 1995a, Pattern recognition techniques for the identification of cloud and cloud systems. *Meteorological Applications*, **2**, 257-271.

PANKIEWICZ G.S., 1995b, A neural network cloud classifier trained with AVHRR data for use on Meteosat imagery. *Meteorological Satellite Data Users' Conference, Proceedings of the EUMETSAT Conference held in Winchester, UK, on 4-8 September 1995*, EUM P17 (Darmstadt: EUMETSAT), pp.393-400.

PEAK, J.E. AND TAG, P.M., 1992, Toward automated interpretation of satellite imagery for Navy shipboard applications. *Bulletin of the American Meteorological Society*, **73**, 995-1008.

PIERCE, C.E., COLLIER, C.G. AND HARDAKER, P.J., 1995, Forecasts of heavy convective rain for use within a flash flood warning system. *National Hydrology Symposium, Proceedings of the 5th BHS Symposium held in Edinburgh, UK, on 4-7 September 1995*, (British Hydrological Society), pp.4.17-4.23.

RUMELHART, D.E., HINTON, G.E. AND WILLIAMS, R.J., 1986, Learning representations by back-propagating errors. *Nature*, **323**, 533-536.

TOVINKERE, V.R., PENALOZA M., LOGAR A., LEE J., WEGER R.C., BERENDES T.A. AND WELCH, R.M., 1993, An intercomparison of artificial intelligence approaches for polar scene classification. *Journal of Geophysical Research*, **98 D3**, 5001-5016.

WELCH, R.M., SENGUPTA, S.K., GOROCH, A.K., RABINDRA, P., RANGARAJ, N. AND NAVAR, M.S., 1992, Polar cloud and surface classification using AVHRR imagery: An intercomparison of methods. *Journal of Applied Meteorology*, **31**, 405-420.

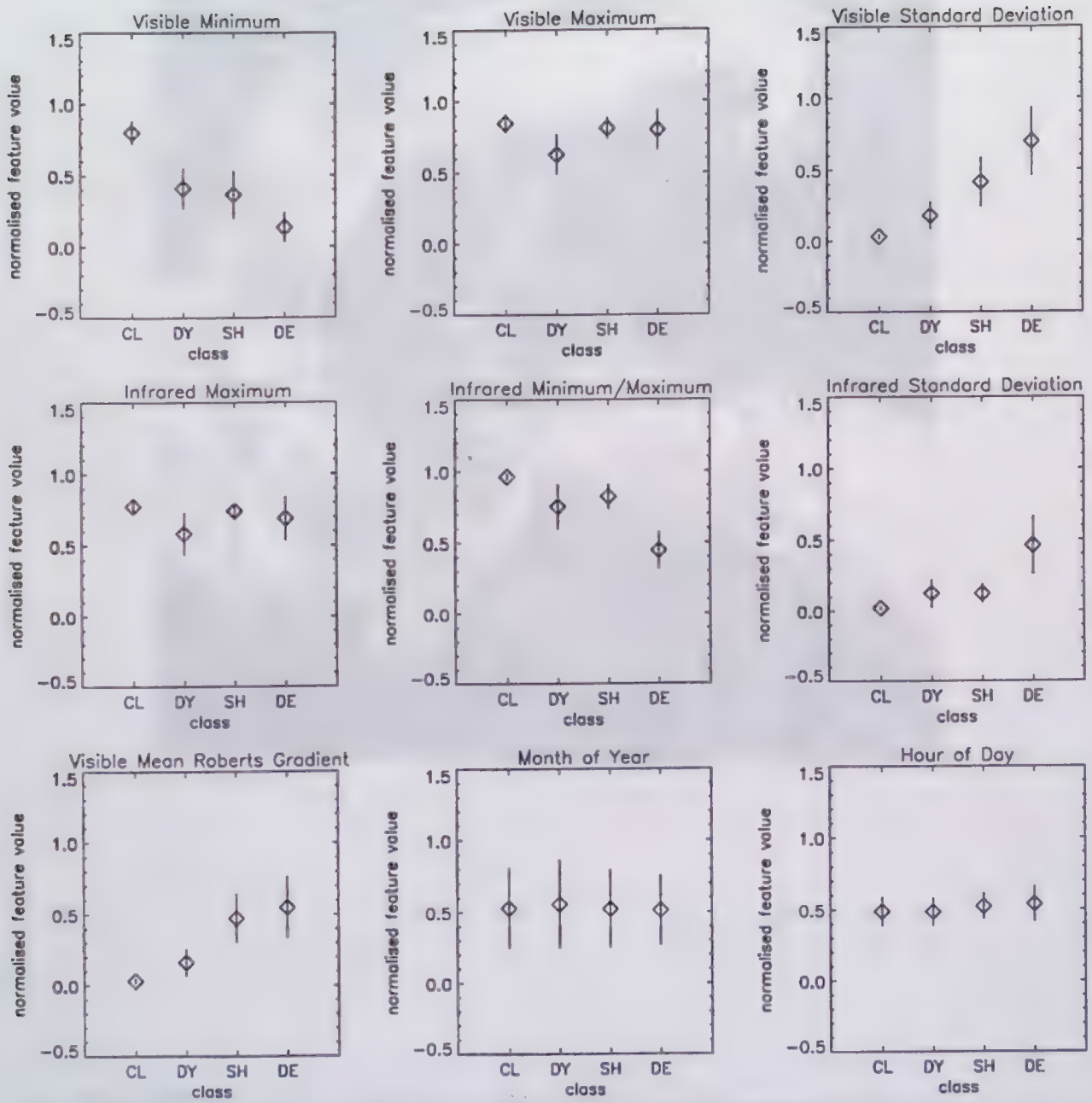



Figure 1: The 9 normalised features used in the GANDOLF neural network. The mean and standard deviation of each feature is shown for all four classes. Note that the original Autosat-2 grey-scale values were used, hence the minimum visible mean for clear sea and land is 0.80, corresponding to an albedo of 20%.

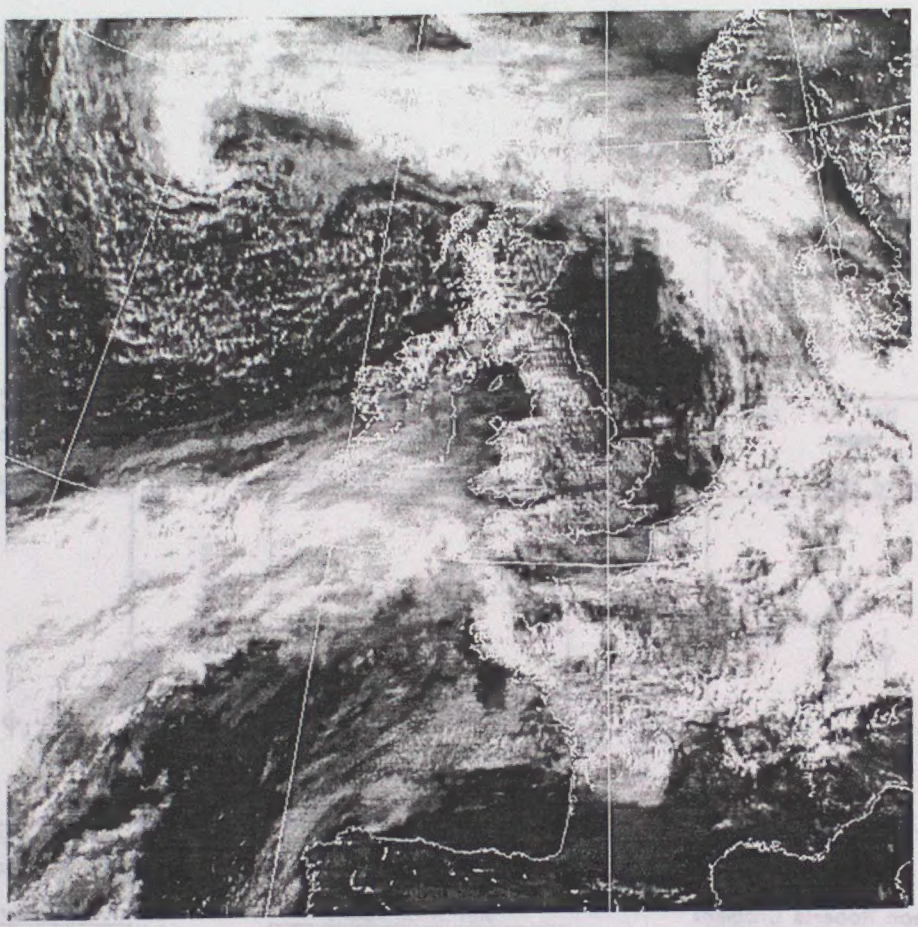
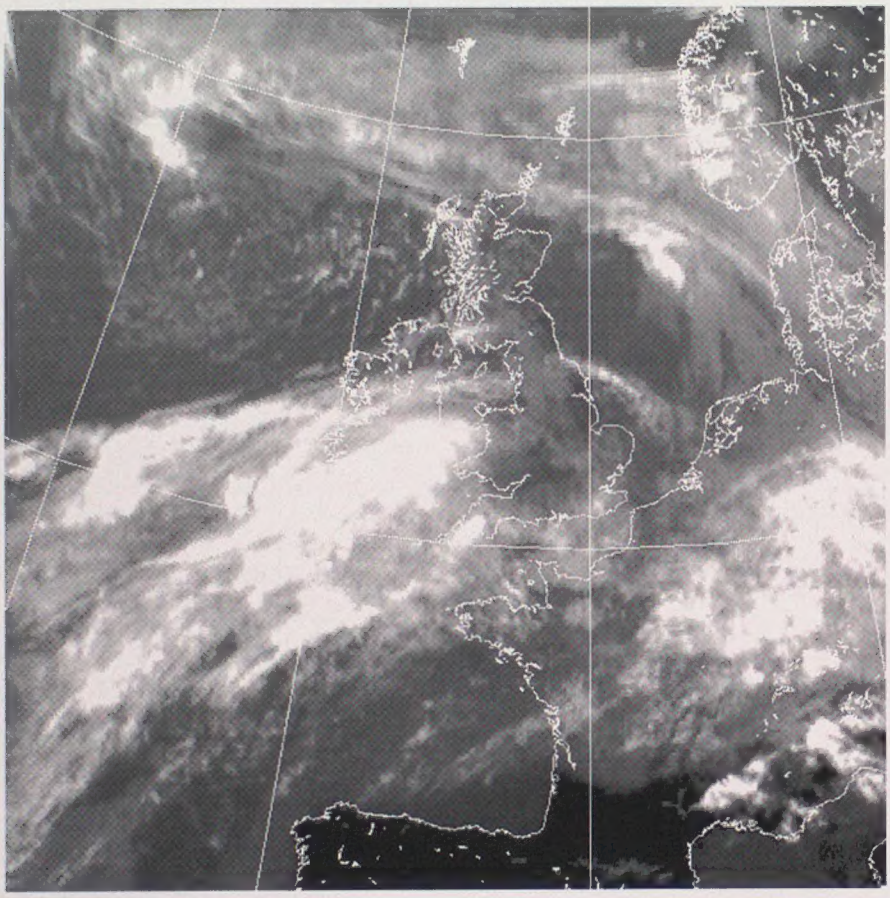


Figure 2: A visible Meteosat image taken on May 23rd 1996 at 10:30 UT with overlying coastline and latitude-longitude grid, covering the Autosat-2 EB area. The grey-scale has been flipped so that cloud appears bright.



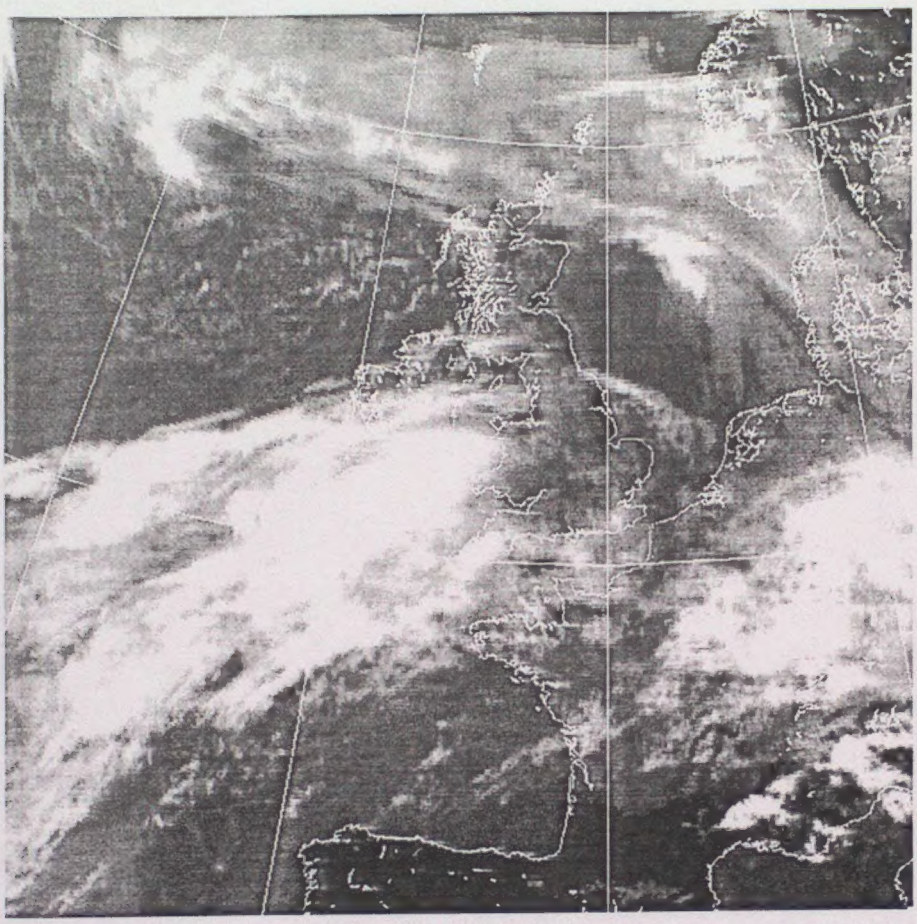


Figure 3: The corresponding infrared Meteosat image to that shown in figure 2. Again, the grey-scale has been flipped so that cold cloud appears bright.

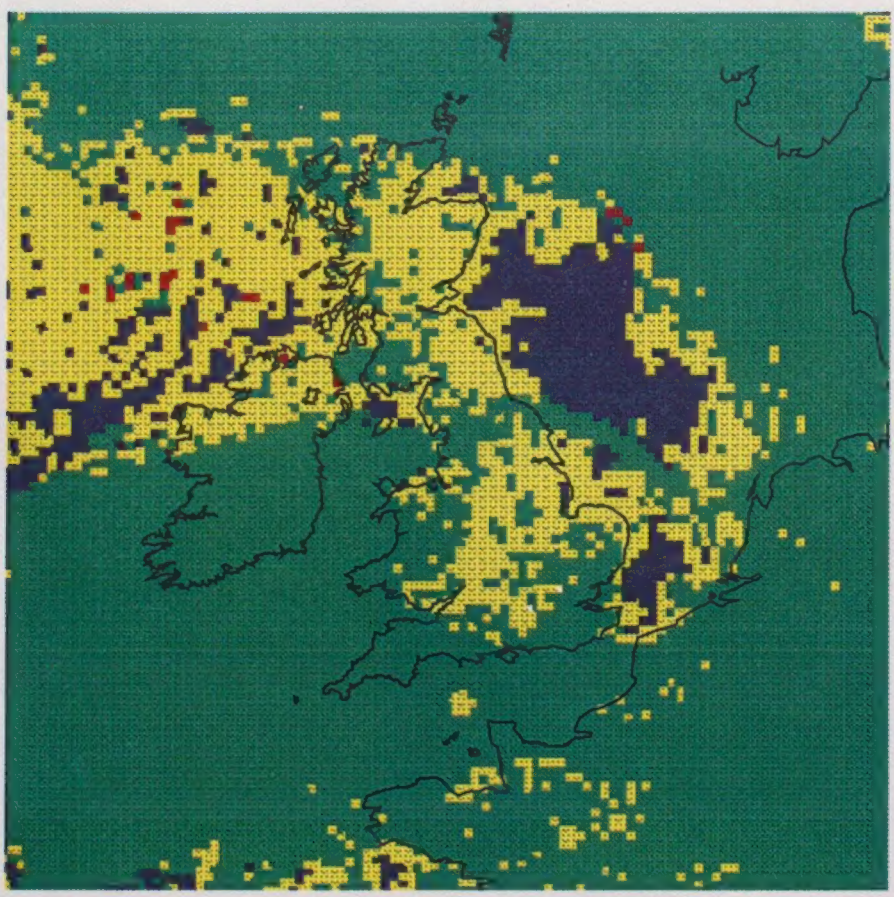


Figure 4: The cloud classified image for May 23rd 1996 at 10:30 UT obtained using the neural network described in the text. The image covers roughly the central two-thirds of the Meteosat images shown in figures 2 and 3, to coincide with the Nimrod area. The following colour code is used: Blue = clear land and sea, green = dynamic cloud, yellow = shallow convection and red = deep convection.